

Scotland's Rural College

## Towards on-farm pig face recognition using convolutional neural networks

Hansen, MF; Smith, Melvyn L; Smith, LN; Salter, MG; Baxter, EM; Farish, M; Grieve, B

*Published in:*  
Computers in Industry

*DOI:*  
[10.1016/j.compind.2018.02.016](https://doi.org/10.1016/j.compind.2018.02.016)

First published: 21/03/2018

*Document Version*  
Peer reviewed version

[Link to publication](#)

### *Citation for pulished version (APA):*

Hansen, MF., Smith, M. L., Smith, LN., Salter, MG., Baxter, EM., Farish, M., & Grieve, B. (2018). Towards on-farm pig face recognition using convolutional neural networks. *Computers in Industry*, 98, 145 - 152.  
<https://doi.org/10.1016/j.compind.2018.02.016>

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Towards on-farm pig face recognition using convolutional neural networks

Mark F. Hansen<sup>\*1</sup>, Melvyn L. Smith<sup>1</sup>, Lyndon N. Smith<sup>1</sup>, Michael G. Salter<sup>2</sup>, Emma M. Baxter<sup>3</sup>, Marianne Farish<sup>3</sup>, and Bruce Grieve<sup>4</sup>

<sup>1</sup>Centre for Machine Vision, BRL, UWE Bristol, Bristol BS16 1QY, UK

<sup>2</sup>AB Agri Ltd. Innovation Way, Peterborough

<sup>3</sup>Animal Behaviour and Welfare, Animal and Veterinary Sciences Research Group, SRUC, West Mains Road, Edinburgh EH9 3JG, UK

<sup>4</sup>School of Electrical and Electronic Engineering, University of Manchester, M13 9PL, UK

## Abstract

Identification of individual livestock such as pigs and cows has become a pressing issue in recent years as intensification practices continue to be adopted and precise objective measurements are required (e.g. weight). Current best practice involves the use of RFID tags which are time-consuming for the farmer and distressing for the animal to fit. To overcome this, non-invasive biometrics are proposed by using the face of the animal. We test this in a farm environment, on 10 individual pigs using three techniques adopted from the human face recognition literature: Fisherfaces, the VGG-Face pre-trained face convolutional neural network (CNN) model and our own CNN model that we train using an artificially augmented data set. Our results show that accurate individual pig recognition is possible with accuracy rates of 96.7% on 1553 images. Class Activated Mapping using Grad-CAM is used to show the regions that our network uses to discriminate between pigs.

**Keywords** Pig face recognition, deep learning, convolutional neural network, biometrics

## 1 Introduction

The need for on farm identification of individual animals has become more pressing in recent years as sustainable intensification has become commonplace, and the ability to monitor inputs to, and outputs of each animal is increasingly desired. The major method of livestock identification is via passive Radio Frequency IDentification (RFID) tags. These low-cost tags are commonly fitted to the animals' ears through piercing – a time consuming and distressing activity for the animal. They also have a limited range (even long range readers state a maximum distance of 120cm) at which they can be activated and read successfully, and multiple tags cannot be read concurrently. Even fitting two tags per pig (to improve the chance of a successful reading) was found to only identify the animal at close range with an accuracy of 88.6% [1]. Common elements in the farm environment can also be detrimental to the antenna's effectiveness. Metal apparatus and tag readers from other equipment (e.g. shedding gates or weigh scales) can reduce the range further and interference can prevent some equipment from functioning at all.

Human face recognition has been an active area of research for at least five decades [2]. From geometric feature matching to holistic methods in the 1990s ([3, 4]), the recent trend of using deep networks has advanced the state-of-the-art to near human level performance [5, 6]. It is commonly used for non-intrusive access control and monitoring/surveillance purposes, and as such represents a potentially useful research area to apply to the problem of pig identification. Although there has been related work to automatically identify behaviours of pigs [7] and feeding/standing of cattle [8, 9, 10], biometrics on cattle [11, 12, 13, 14], sheep [15] and canines [16] showing promising results, to date there has been very little research into using a pig face as a biometric, although [17]

---

<sup>\*</sup>mark.hansen@uwe.ac.uk

show some preliminary results of applying the Eigenfaces technique to pigs and achieve a recognition performance of 77% on 10 pigs using the full manually cropped face. They reported better results for smaller regions (i.e. the nose, or the eyes), but this relied on further manual segmentation of the regions so is not very applicable to an on-farm system. They also only collected 16 sequential image frames per pig, so the generalisation of such a system to different environmental conditions when imaging the same pigs is unknown.

This paper presents the results of three face recognition methods applied to a dataset of pig faces that have been captured on a farm under natural conditions: Fisherfaces [4], transfer learning using the pre-trained VGG-Face model [6] and our own convolutional neural network which has been trained using our own dataset captured using an off the shelf web camera at the drinker in a pen. This represents a machine vision application in a challenging, poorly structured environment and even though the pigs are technically under cover, ie in a shed, the subjects position and pose as well as other aspects, such as the lighting, expression, contamination from dirt, etc., are relatively uncontrolled and highly variable. We demonstrate the efficacy of the system on recognising unconstrained, and un-preprocessed images of pig faces and present an analysis of those features and activation areas which our system has learned in response to training.

The rest of the paper is organised as follows: Section 2 outlines the data collection methodology, the video preprocessing approach to remove very similar frames, and the implementation details. Section 3 gives the background to the chosen approaches before the results being given in Section 4. A discussion follows which puts these results into context together with limitations and suggestions for future work.

## 2 Material and methods

This section describes the data capture, data cleaning and implementation details. An overview of the processing pipeline can be seen in Figure 1.

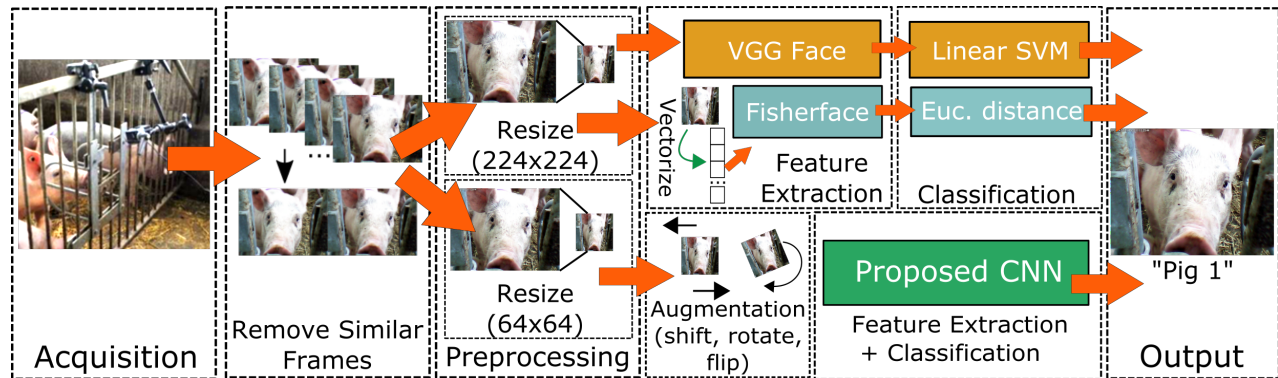


Figure 1: Processing pipeline showing the acquisition, pre-processing steps, feature-extraction and classification for the three methods used in this paper for pig face recognition.

### 2.1 Data collection

The pigs were Large White x Landrace x Hampshire breed, approximately four months old and housed at SRUC's research farm (Midlothian, Scotland). The pigs were filmed using a Sogatel USB2.0 webcam, with VGA resolution (640x480 px) at 30 frames per second. The camera was connected to a Dell Precision laptop running "iSpy Connect" software to allow motion-detection capture of the pigs each time they voluntarily approached the drinker. The camera was positioned behind the drinking nipple as shown in Fig. 2. A Manfrotto universal clamp and articulated arm were used to mount the camera to the pen frame at sufficient distance to ensure it was out of reach but close enough that the bars did not obscure the pigs' face as they drank. Access to the drinker was altered slightly through addition of shoulder bars which help to keep the pig face-on to the camera and other pigs from being in frame. No other changes were made to the drinker and the experiment was approved by the SRUC's Animal Ethics Committee.



Figure 2: Photographs showing the modified drinking nipple (left) and the arrangement of the webcam behind it (right).

Data of 10 pigs were collected in 2 sessions (31/03/2017 and 03/04/2017). The camera was left running unattended and manually labelled afterwards in order to create the training and test data required. As can be seen in the images, the pigs have been spray-painted to aid manual identification, and this is not required by the automated system itself. It should also be noted that some care was taken to mount the camera in such a way as to ensure that direct sunlight did not fall on the pigs' faces at the drinker as this led to saturated images. Other than this, natural variation in lighting levels was handled automatically by the camera. Examples of the pigs can be seen in Fig. 3.



Figure 3: Photographs of the ten pigs used in this study.

## 2.2 Data cleaning

To avoid the shortcoming noted in [17] regarding low variance between consecutive frames, the structural-similarity index measure (SSIM) [18] is employed to measure similarity between images. This helps to avoid very similar (near identical) data in train and test data partitions. This approach attempts to be more similar to human



perception than the alternative and commonly used mean square error (MSE) measure in reporting similarity between images. It takes into account the variance, covariance and mean intensities within two images,  $x$  and  $y$  as shown in Eqn. 1 where  $\mu_x, \mu_y$  are the average,  $\sigma_x, \sigma_y$  are the variance and  $\sigma_{xy}$  the covariance of images  $x$  and  $y$ .  $c_1, c_2$  are constants to avoid instability when the denominator is close to zero.

$$SSIM_{x,y} = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (1)$$

Each image is compared to subsequent images until a sufficient difference (an empirically determined threshold) is found (Fig. 4).

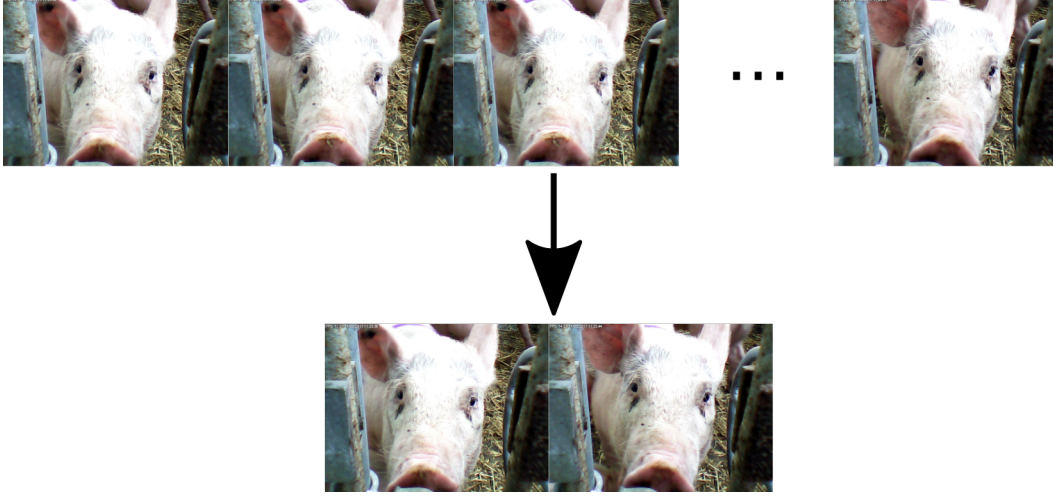


Figure 4: Examples of consecutive raw captures, and those frames which are deemed sufficiently different using the SSIM measure. In this instance 31 images captured over 8s have been reduced to just two which at first glance look identical but closer inspection reveal fairly apparent differences e.g. the snout.

The number of raw images for each pig compared with only dissimilar images can be seen in Table 1.

Pig	Raw	Dissimilar
1	659	153
2	492	58
3	596	66
4	691	215
5	489	109
6	418	122
7	440	247
8	409	121
9	730	319
10	529	143
Total	5453	1553

Table 1: The number of images for each pig before and after discarding consecutive images which are deemed too similar to one another. Approximately 70% of images are discarded.

## 2.3 Implementation

The convolutional neural network was written using available libraries in Python3.5 (Keras2.0, scikit-learn0.18.2, tensorflow-gpu-1.1.0). All code has been run on a desktop computer with an Intel i5 processor, 16GB RAM using Windows7 64-bit, and an NVidia Titan-X 12GB graphics card with Maxwell Architecture. Training our network

over 500 epochs takes approximately five minutes, and results of the model from the best performing epoch on the test data is used.

The pretrained VGG-face model was loaded using `matconvnet-1.0-beta13` (a Matlab Toolbox for CNNs), and Matlab's default Linear SVM implementation was used in Matlab 2017a. This application was also used for the Fisherface implementation and uses  $N - M$  coefficients for PCA, to generate  $M - 1$  LDA coefficients as per the original formulation [4], where  $N$  is the number of observations and  $M$  is the number of classes.

### 3 Theory

This section presents the algorithms and neural network architectures used for the identification of individual animals. Due to the abundance of work in human face recognition [19], it is logical to apply well proven methods from that area to explore the feasibility of pig face recognition. In particular we use two approaches: a benchmark method known as Fisherfaces and a deep learning approach which we split into two formulations: a deep pre-trained convolutional network model that we classify using a Support Vector Machine (SVM) and our own convolutional network which uses a fully connected layer for classification.

#### 3.1 Fisherfaces

Fisherfaces uses a combination of Principal Component Analysis (PCA) with Fisher's Linear Discriminant (FLD). FLD is named after Robert Fisher who developed the technique for taxonomic classification [20]. The key aspect to this technique is that it uses *labelled* data and seeks to minimise intra-class scatter and maximise inter-class scatter. The concept of clusters is important in FLD, and ideally the cluster for a given label (or class) is compact (small intra-class scatter) and distant from other clusters (large inter-class scatter). This lends itself well to face recognition as faces are labelled as being that of a certain person.

A major assumption here is that the intra-class scatter matrix is non-singular. However, when the number of variables is far greater than the number of classes – which is the case when used for face recognition (i.e. the number of pixels is far greater than the number of identities), the likelihood of the matrix being singular is extremely high. To overcome this, Belhumeur *et al.* use PCA on the image set as a means of reducing the dimensionality first before projecting into FLD space. It is this additional step which allows them to identify their approach as *Fisherfaces*.

In their paper, Belhumeur *et al.* compare performance of PCA (Eigenfaces) and Fisherfaces and report far better results for Fisherfaces under varying illumination and expression (error rates half that of any other method and a third of Eigenface).

Although this work is over 20 years old, it is still prominently used as a benchmark method due to its effectiveness, hence its inclusion in this paper.

#### 3.2 Deep learning and convolutional neural networks

Neural networks have been an active area of research for many decades due to their theoretical ability to model any relationship between input and output, linear or non-linear, provided with sufficient data from which to generalise. As far back as 1980, Fukushima [21] proposed an architecture based on the human visual receptive fields that he named the Neocognitron. It described alternating layers that convolved and sub-sampled an input image. This architecture inspired the development of LeNet-5 [22], a 7-layer convolutional neural network (CNN) which recognised hand written digits. However, due to the large numbers of trainable parameters needed, it could only operate on 32x32px images. With the growth in low cost, powerful graphics cards, it is now possible to design and train far deeper networks on desktop machines. For example an NVidia Corporation Titan-X card (similar to that used in this study) contains 3072 computing cores and 12GB of onboard RAM. In recent years deep learning and convolutional networks have been applied in a wide range of both research areas and also in industrial settings [5, 23, 24] with the first reported use for face recognition by Lawrence *et al.* [25].

Deep learning, and CNN approaches in particular, combine both the feature extraction and classification stages of classical pattern recognition, by propagating the training of a fully connected classifier back through the convolutional layers in order to select the best features. A popular publicly available model that offers close to state-of-the-art performance on the Labelled Faces in the Wild dataset [26] is called VGG-Face which is based

on the VGG-Very-Deep-16 CNN described in [6] and contains 37 layers. The model has been trained on 2.5 million images of 2400 people. In doing so it will have learned discriminative features of the faces. We propose to use this pre-trained model to see whether we can leverage the power it shows in humans on pig faces. In order to do this we need to either retrain the classification layers at the end of the network with the new classes, or feed the 4096 features that are output from the final convolutional layer into a traditional classifier. We were unable to get a new classification layer to converge using the pig photographs (presumably because there is an insufficient number given the depth of the network), and so settled for using a Support Vector Machine (SVM) to discriminate between the 4096 features to identify individuals.

To train a network of this depth not only takes a great deal of computing power, but also a very large amount of training data. In this study we have a limited amount of data (on average 150 images per pig), and so are limited in how deep we can make our own network. In order to boost the amount of data we perform small manipulations to increase the variance and also improve the robustness of the network (shifting the images by up to 64 pixels in the x/y plane, flipping the images horizontally and rotating up to 30 degrees). The colour jpg images are 640x480 pixels in size and these are resized to 64x64 pixels (this was determined empirically to give an acceptable compromise of good performance vs speed of processing), before being fed into the first layer of the network. The target for each image is the identity of the pig (1 to 10), and these are converted to a 'binary' representation of 10 output nodes, where a '1' in the relevant index corresponds to that identity e.g. '0100000000' corresponds to Pig 2.

In both networks, the data is split 60:40 into train and test partitions, and the results in Section 4 report performance on the test partition.

The network we train consists of six convolutional layers, with alternating dropout and max-pooling layers in between. The purpose of these is to aid convergence by preventing local minima, and providing scale/location invariance. The classification layer consists of three fully connected layers with the final layer containing 10 outputs – corresponding to each pig. The architecture is shown in Fig.5.

## 4 Results

Using the dissimilar dataset, the results on the test partition using the Fisherface, VGG-face pre-trained model and our own network can be seen in Table 2. Accuracy is given as the number of correctly identified images as a percentage of the whole dataset. The False Positive (FP) rate is the number of pigs that are incorrectly identified as a particular pig as a percentage of total identifications of that pig ie how many times an imposter is identified as a certain pig e.g. for pig 1, if pig 4 is identified as pig 1, then this is a false positive. Likewise, the False Negative (FN) rate is the number of times a particular pig is wrongly identified as another pig as a percentage of total identifications of that pig e.g. for pig 1, if pig 1 is identified as pig 4, then this would be a false negative. We essentially replicate the previous findings of [17] using Fisherfaces instead of Eigenfaces and achieve an accuracy of 78.4%. This is very similar to their reported accuracy of 77% on whole faces. The human face recognition literature tends to report better results for Fisherfaces over Eigenfaces and one reason we are not seeing a marked improvement may be due to their results reporting on very similar testing and training images, together with the manual cropping of the pig faces which they performed.

However, the Fisherface results are convincingly outperformed by both deep learning approaches. Extremely high recognition results are seen using the CNNs with our own performing the best (just under 97%). The entire test set of 622 images can be identified in less than one second (0.002s per image) once the network has been trained, meaning that the system can be run on a standard computer to identify animals in real time.

The FP rates shown in Table 2 show the number of false identifications that take place. These are important metrics in terms of how practical the system might be in real-world use. The Fisherface and VGG-Face+SVM approaches show the highest FP and FN rates with over 1 in 4 pigs potentially being misclassified in the case of Fisherfaces. This rate is very likely too high for on farm use. Our proposed CNN approach does better at 5.8%, meaning that just over 1 in 17 might be misidentified. However, much of this error comes from the confusion between pigs 2 and 3 which have high false positive rates of 11% and 25% respectively.

It is very interesting that the VGG-Face pre-trained model performs as well as it does given that it has only ever been trained on human faces. This indicates that many of the same features that the network has learned to be useful for discriminating human faces are also useful for discriminating pig faces. This hints toward how a trained network for faces in one species may be transferable to other species, at least in so far as using it to

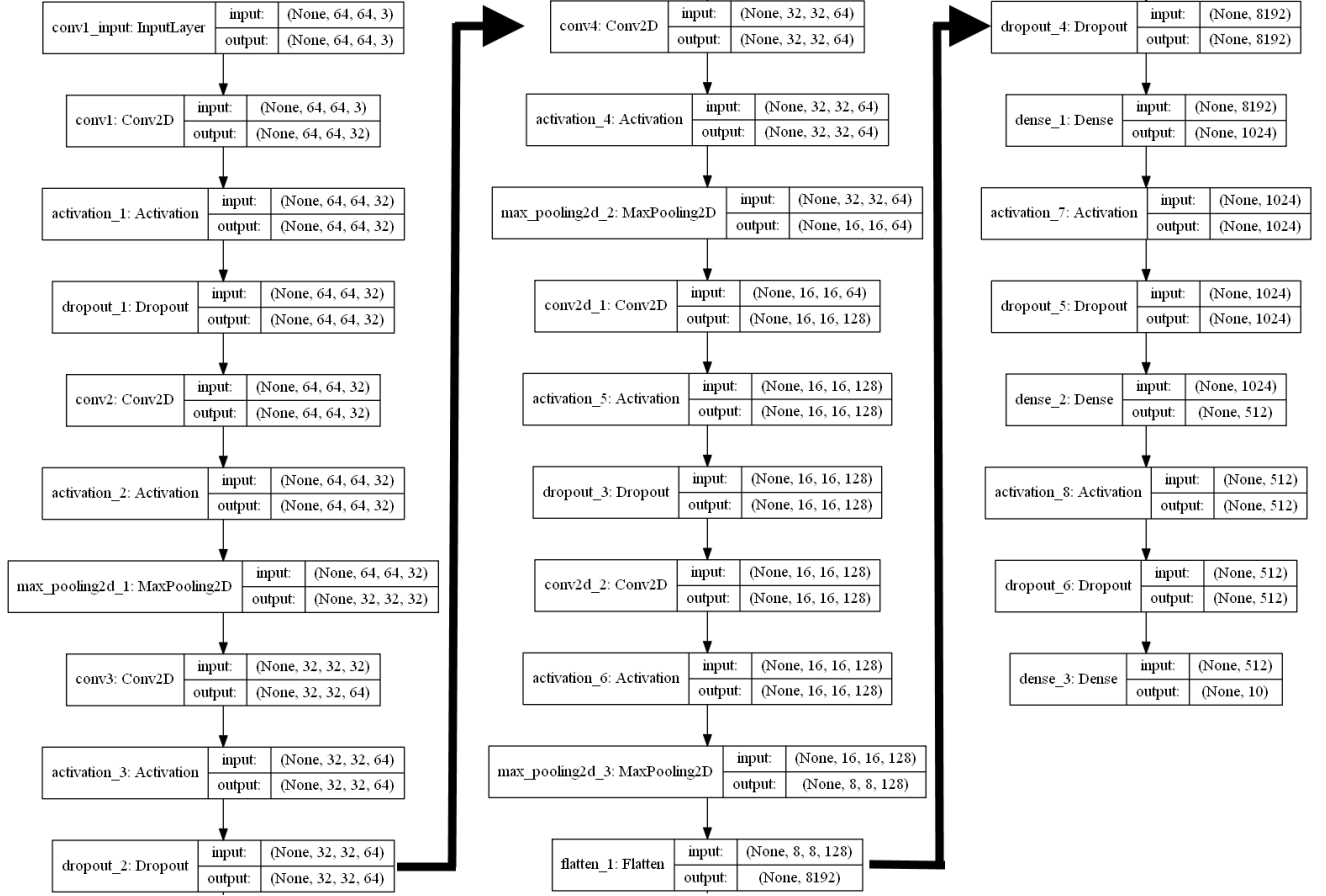


Figure 5: Architecture of our CNN consisting of six convolutional layers with alternating dropout and max-pooling layers. Classification occurs in the final three fully connected layers.

bootstrap a new model for a new species using a reduced training set. This bootstrapping approach will be the subject of future work.

Method	Accuracy (%)	False Positive (%)	False Negative (%)
Fisherface	78.4	26.3	29.7
VGG-Face + SVM	91.0	10.3	13.9
Our CNN	96.7	5.8	5.7

Table 2: Summary results on the test partition for three face recognition approaches on 10 pig faces using 931 training samples and 622 test samples (60:40 train:test split).

Fig. 6 shows the normalised confusion matrix for the results of our CNN. The pig which is misidentified the most is Pig 2 (67% accuracy), which is confused for Pig 3 in 21% of cases and Pig 6 in 12% of cases. Interestingly, referring to Fig. 3, these three pigs do not have pigmentation/markings on their faces (unlike the other seven pigs). What is uncertain is why there should be asymmetry in the confusion matrix i.e. Pigs 3 and 6 have recognition accuracies of 96% and 94% respectively, but Pig 2 only has 67%. The number of samples for Pig 2 is only slightly less than for Pig 3 (58 vs 66), although these two are the only two with less than 100 samples, so this could be a contributory factor. Indeed as discussed above, the FP rate for these two pigs is also far higher than the others; another indication that the number of samples is probably too low – something which will be addressed in further studies by increasing the size of the dataset.

The Receiver Operating Characteristic (ROC) plot shown in Fig. 7 demonstrate the robustness of the system with extremely high areas under each curve indicating high accuracy with very low false positive rates. As expected,

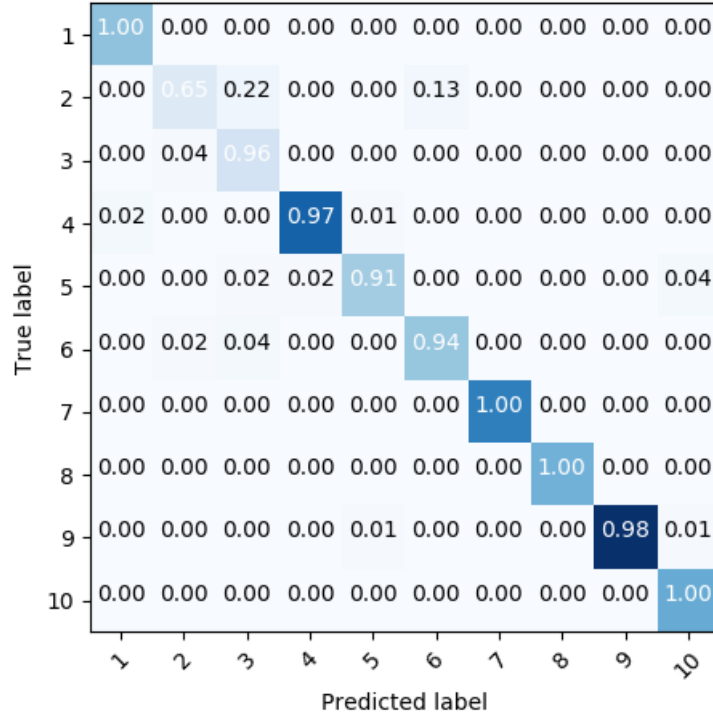


Figure 6: Normalised confusion matrix of our CNN. Note the major confusion occurs between pigs 2 and 3, but not vice versa.

Pig 2 generates the worst performance, with four of the pigs (1, 7, 8 and 10) being correctly identified in every instance.

## 5 Discussion

In summary, the results indicate that it is possible to accurately recognise individual pigs non-invasively from a relatively unconstrained scene (i.e. the pig can present their faces in many poses). On average, using our own trained CNN we achieve an accuracy of 96.7% on 1553 images of 10 pigs. This outperforms a standard face recognition technique frequently used in automated human face recognition known as Fisherfaces and also a pre-trained human face recognition CNN, VGG-face.

A frequently cited downside of using neural networks is that understanding exactly what they are modelling is very difficult. This is further compounded using deep networks. However, this has begun attracting a great deal of research interest, especially for CNNs to ensure that the "attention" of the network is focussed on real, discriminative features of the animal itself, rather than other parts of the image which may happen to contain discriminative information (e.g. a class label, a time stamp imprinted onto the image or even the spray-marked region of the pig). In [27], Selvaraju *et al.* present a method of producing a coarse localization map for a given class that the network has been trained on (Gradient-weighted Class Activated Mapping, Grad-CAM). Essentially this shows what regions of an input image are activating the network for a given class. We implement this approach and show some representative examples in Fig. 8

By examining Fig. 8 it is possible to see that the network learns to discriminate the pigs mainly from three regions: the snout itself and the region above it containing wrinkles, the top of the head (where the markings are most prevalent) and, to a lesser extent, the eye regions. Unfortunately it does not appear to give any insight into performance variations i.e. it does not tell us anything obvious about why the recognition performance on Pig 2 is poor, nor why the performance is at 100% for pigs 1, 7, 8 and 10.

One of the main aims of this work is to assess the feasibility of on-farm livestock recognition systems. Although



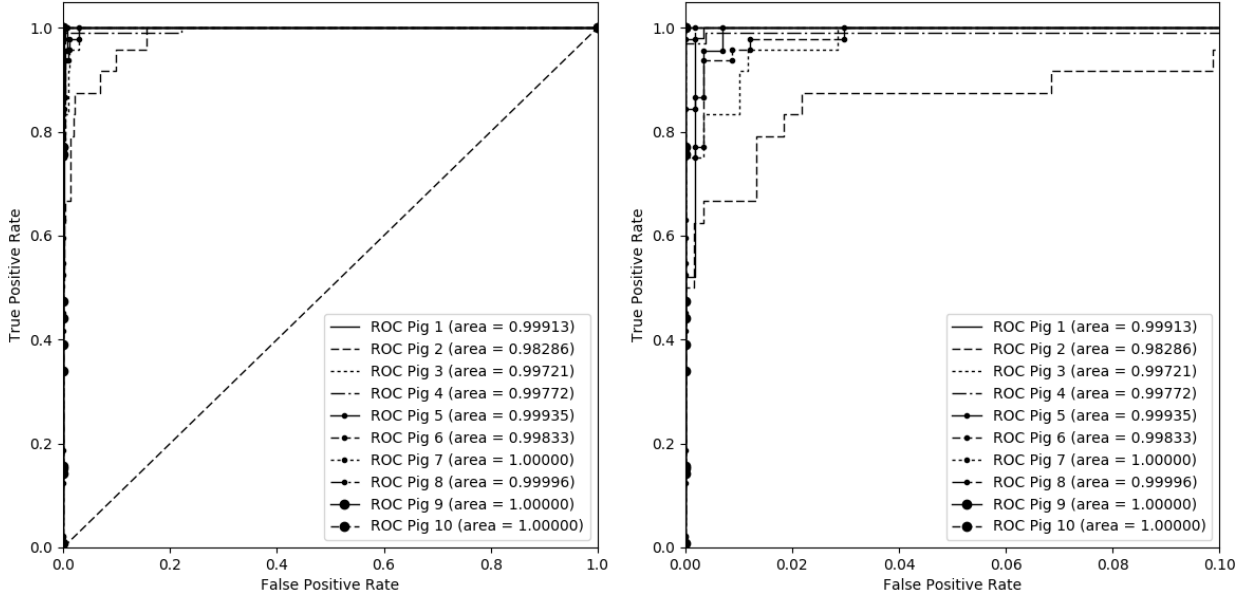


Figure 7: Receiver operating characteristic for the recognition of each pig (and the enlarged region for  $FP=0,0.1$ ). At a 6% false positive rate, all pigs are recognised with over 90% accuracy. Eight of the pigs, achieve an accuracy of 95% at a false positive rate of 0.8%

the number of animals in this study (10) is relatively small and the images have not been taken over a long enough period to see any large changes in the animal, we have shown that with only minor changes to the arrangement of the drinker, and the addition of a webcam, extremely accurate recognition performance is possible. A drawback with the current arrangement is that the camera is mounted some distance behind the drinker, which in practice would locate it in an adjacent pen. The reason for choosing this location was to maximise the opportunity of getting frontal faces of the pigs while they drank. From the results presented in this paper, and specifically the regions which the network uses to identify the individuals, it is reasonable to hypothesise that an entirely overhead view is likely to provide sufficiently good data without the need to put the camera in an adjacent pen. This could be done in conjunction with other technology such as a 3D camera which would be capable of estimating the volume and hence weight of the animal.

While this project specifically looked at pigs there is reason to believe that such an approach may be used for other common livestock breeds e.g. cows and sheep, where precision non-invasive monitoring is required. While no effort was made to ensure that the pigs' face were not dirty in this study, none appeared to become overly muddy. However this need not necessarily become too much of an issue (within reason); as long as the training data shows sufficient variance of the animals with dirty faces, reliable descriptors will be extracted that do not rely on these features.

Other factors that may affect the recognition over the duration the animal is housed are: how the animals' face ages, how big it grows and in particular relevance to pigs, the presence of tear stains beneath the eyes which may also be used as a welfare indicator [28]. These can have a marked effect on the appearance of the animal, and future work will look to quantify and mitigate the effects. In terms of an on-farm solution, one could retrain the network on a daily basis, using recently taken images. Assuming that any changes are slow (e.g. which seems reasonable in the case of growth/ageing) the animal will be correctly recognised and these incremental changes will become incorporated into the network's model in an unsupervised manner. The effects from more instantaneous changes to an animals' appearance due to dirt are unknown but will be the focus of further work.

It would also be beneficial to measure the effectiveness on a large number of pigs that do not have any markings and see whether the system can still learn to distinguish them based on different features. We might expect the VGG-face model to perform well given data of this sort as the data are arguably more similar to human faces. If this work shows that the proposed system is not scalable to large numbers of pigs, a potential solution

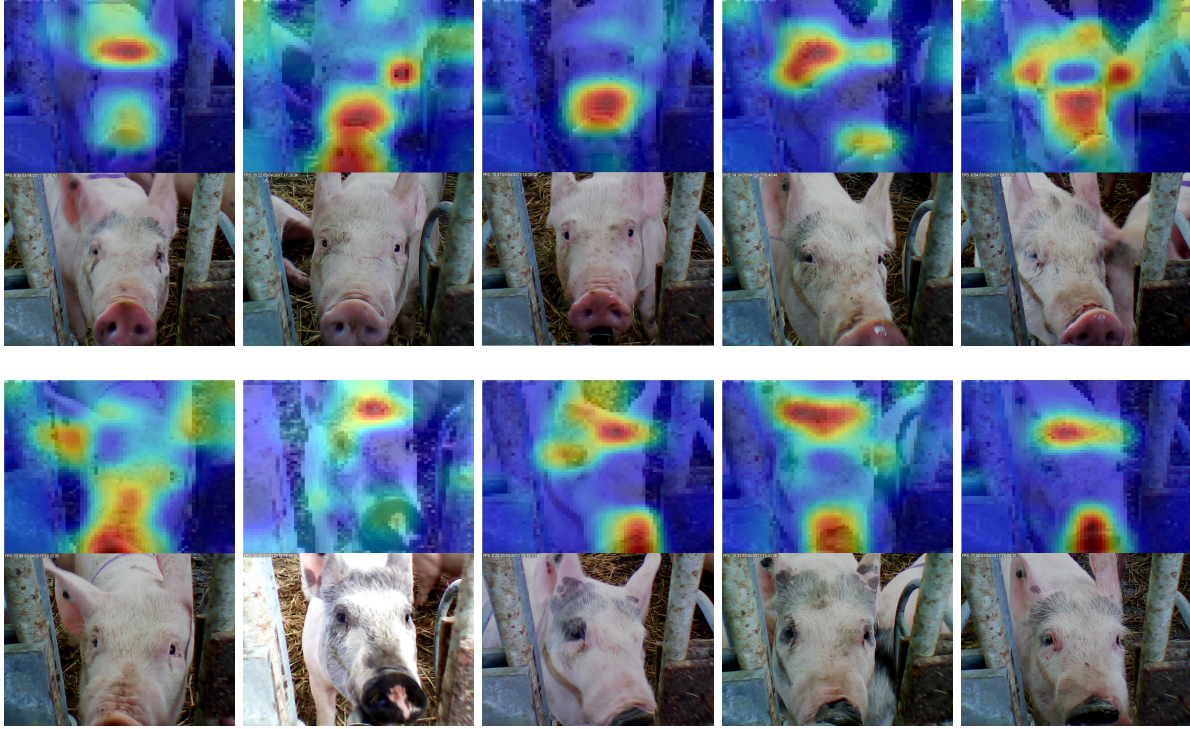


Figure 8: Class activated maps overlaid on input images together with the original photographs. This gives us confidence that the network is not learning a discriminative feature based on something other than the pig’s face (e.g. the time stamp). Pigs are shown in the same as order as Fig. 3, top row shows pigs 1-5, and the bottom row shows pigs 6-10.

would be to consider using the system at a "pen-level" rather than across the entire farm. This would ensure fewer pigs for each system, but still provide a feasible method of identification of individual pigs across the farm. Future work will investigate these effects, and test the scalability limits of the system.

## 6 Conclusion

We have presented a non-invasive imaging system capable of recognising individual pigs from their faces at a minimally adjusted drinker in their pen. The system uses data from an unconstrained, commercial farm environment where the animals’ position and pose as well as other aspects, such as the lighting and dirt, are relatively uncontrolled and highly variable. Once trained it can operate in real time and accurately identify pigs with a high accuracy (96.7%). Although our limited dataset consists of 1553 images of 10 pigs, the excellent results presented here demonstrate the potential of our approach, and further work will look to use a more convenient overhead viewpoint and investigate the effects of more changeable aspects of the pig’s appearance (e.g. ageing, dirt, tear staining). The study has implications for intensive livestock practices globally, allowing identification of animals without the need for RFID tags, for the purposes of welfare and growth monitoring.

## 7 Acknowledgements

This research has been funded by Natural Environment Research Council, UK (Grant Reference: NE/P007945/1) as part of the Sustainable Agriculture Research & Innovation Club in collaboration with AB Agri. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Maxwell Titan X GPU used for this research.

## References

- [1] J. Maselyne, W. Saeys, B. De Ketelaere, K. Mertens, J. Vangeyte, E. F. Hessel, S. Millet, and A. Van Nuffel. Validation of a High Frequency Radio Frequency Identification (HF RFID) system for registering feeding patterns of growing-finishing pigs. *Computers and Electronics in Agriculture*, 102(Supplement C):10–18, March 2014.
- [2] W. W. Bledsoe. Man-machine facial recognition. *Panoramic Research Inc., Palo Alto, CA*, 1966.
- [3] M. Turk and A. Pentland. Eigenfaces for Recognition. *J. Cognitive Neurosci.*, 3(1):71–86, 1991.
- [4] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Trans. Patt. Anal. Mach. Intell.*, 19(7):711–720, 1997.
- [5] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf. Deepface: Closing the gap to human-level performance in face verification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1701–1708, 2014.
- [6] O. M. Parkhi, A. Vedaldi, and A. Zisserman. Deep Face Recognition. In *Proceedings of the British Machine Vision Conference (BMVC)*, volume 1, pages 41.1–41.12, 2015.
- [7] P. Ahrendt, T. Gregersen, and H. Karstoft. Development of a real-time computer vision system for tracking loose-housed pigs. *Computers and Electronics in Agriculture*, 76(2):169–174, May 2011.
- [8] S. M. C. Porto, C. Arcidiacono, U. Anguzza, and G. Cascone. A computer vision-based system for the automatic detection of lying behaviour of dairy cows in free-stall barns. *Biosystems Engineering*, 115(2):184–194, June 2013.
- [9] S.M.C. Porto, C. Arcidiacono, U. Anguzza, and G. Cascone. The automatic detection of dairy cow feeding and standing behaviours in free-stall barns by a computer vision-based system. *Biosystems Engineering*, 133(Supplement C):46–55, May 2015.
- [10] C. Arcidiacono, S. M. C. Porto, M. Mancino, and G. Cascone. Development of a threshold-based classifier for real-time recognition of cow feeding and standing behavioural activities from accelerometer data. *Computers and Electronics in Agriculture*, 134(Supplement C):124–134, March 2017.
- [11] T. Gaber, A. Tharwat, A. E. Hassanien, and V. Snasel. Biometric cattle identification approach based on Weber’s Local Descriptor and AdaBoost classifier. *Computers and Electronics in Agriculture*, 122(Supplement C):55–66, March 2016.
- [12] S Kumar, S.K. Singh, and A.K. Singh. Muzzle point pattern based techniques for individual cattle identification. *IET Image Processing*, 11:805–814, October 2017.
- [13] A. Tharwat, T. Gaber, and A.E. Hassanien. Cattle Identification Based on Muzzle Images Using Gabor Features and SVM Classifier. In *Advanced Machine Learning Technologies and Applications*, Communications in Computer and Information Science, pages 236–247. Springer, Cham, November 2014.
- [14] A. Allen, B. Golden, M. Taylor, D. Patterson, D. Henriksen, and R. Skuce. Evaluation of retinal imaging technology for the biometric identification of bovine animals in Northern Ireland. *Livestock Science*, 116(1):42–52, July 2008.
- [15] U. Gonzales Barron, G. Corkery, B. Barry, F. Butler, K. McDonnell, and S. Ward. Assessment of retinal recognition technology as a biometric method for sheep identification. *Computers and Electronics in Agriculture*, 60(2):156–166, March 2008.
- [16] S. Kumar and S.K. Singh. Monitoring of pet animal in smart cities using animal biometrics. *Future Generation Computer Systems*, December 2016.
- [17] N. Wada, M. Shinya, and M. Shiraishi. Pig Face Recognition Using Eigenspace Method. *ITE Transactions on Media Technology and Applications*, 1(4):328–332, 2013.

- [18] Z Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.
- [19] W. Zhao and R. Chellappa. *Face Processing: Advanced Modeling and Methods*. Academic Press, 2006.
- [20] R. A. Fisher. The use of multiple measures in taxonomic problems. *Annals of Eugenics*, 7:179–188, 1936.
- [21] K. Fukushima. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36(4):193–202, 1980.
- [22] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [23] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–9, 2015.
- [24] K He, X. Zhang, S. Ren, and J Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.
- [25] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back. Face recognition: A convolutional neural-network approach. *IEEE Trans. Neural Networ.*, 8(1):98–113, 1997.
- [26] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. *University of Massachusetts, Amherst, Technical Report 07*, 49:1, 2007.
- [27] R.R. Selvaraju, A. Das, R. Vedantam, M. Cogswell, D. Parikh, and D. Batra. Grad-CAM: Why did you say that? *arXiv preprint arXiv:1611.07450*, 2016.
- [28] H. Telkänranta, J. N. Marchant-Forde, and A. Valros. Tear staining in pigs: A potential tool for welfare assessment on commercial farms. *animal*, 10(2):318–325, February 2016.